MICHAEL G. O'NEILL
ATTORNEY AT LAW

By ECF

November 17, 2023

Hon. Vera M. Scanlon
United States Magistrate Judge
225 Cadman Plaza East
Brooklyn, NY 11201

    Re: Coley v. Local 15
       21 CV 5892

Dear Judge Scanlon:

   I am writing to request that the Court reopen discovery for the following purposes:

   a. to require defendant to produce the queries used to produce data in this action and in the related action, Langford et al. v. Local 15 et al., 18-cv-7041 pending in this court (the "Langford Action") and

   b. to require defendant to produce a witness to testify with respect to the above queries.

   c. to require defendant to produce relevant records omitted from production due to incorrectly constructed queries.

 In addition, I am requesting that the Court reconsider its previous ruling denying plaintiff's application to require the defendant to provide year of birth data for its members.

Background

   Defendant is a construction union and plaintiff is a member of that union. Plaintiff alleges that the defendant controls the assignment of jobs to its members and, in doing so, discriminates on the basis of age and race. Defendant maintains several databases in which it keeps track of job referrals and member employment. Statistical analysis of that database is a critical material part of plaintiff's case.

   On May 19, 2022, plaintiff propounded written interrogatories and a request for the production of documents, a copy of which is annexed hereto as Exhibit 1. This demand called for a "dump" of defendant's databases for the relevant time period. Defendant began producing documents before serving a written response to plaintiff's demand. Indeed, by October 22, 2022, defendant had still not served a written response, and I reminded counsel for defendant

of that fact as evidenced by Exhibit 2. Nevertheless, defendant never served a written response to plaintiff's demand. Instead, on June 28, 2023, well after the close of fact discovery, counsel for defendant sent an email in which she acknowledged that defendant had never served a response and had assumed that defendants' responses in Langford "were just as good" for use in Coley. A copy of that email is annexed hereto as Exhibit 3.

Because of this, it is necessary to examine production of the database in Langford to understand plaintiff's expectation of what was being produced in this case. Defendants' response to plaintiffs' demand in Langford is annexed hereto as Exhibit 4. Subsequently, a dispute as to the production of the database in Langford arose, which is evidenced by the email chain annexed hereto as Exhibit 5. As those emails show, defendant ultimately agreed as follows:

> Based on our review of the Magistrate Judge's ruling last month in Sampson v. IUOE Local 14, Local 15 has reconsidered its position and will produce the daily referral database and stamp data providing hours worked for each employee from 2013-2019 subject to limitations on withholding some personal/confidential information such as social security number,

This signaled that production of the database would be via a subtractive process, i.e., the entire database would be produced with certain specified information subtracted from it. This, therefore, is what plaintiff reasonably believed was being produced in this case. That understanding was reinforced by an April 20, 2021 email from counsel for defendant, attached as Exhibit 7 and discussed below.

Defendant reveals production of its database in this case was not the same as in Langford.

On October 17, 2023, I learned from counsel for defendant that "the database that was provided to you in September of 2022 was not produced using the same parameters as had been used in Langford/Sosa." Counsel proposed that the premotion proceedings be stayed for at least 30 days during which time "we could all review the updated production and then re-assess the schedule going forward." A copy of counsel for defendant's October 17, 2023 email is annexed hereto as Exhibit 6. Although the nature, extent or cause of the problem was not clear from Counsel's email, I did agree that proceedings should be put on hold pending a resolution of any issues arising from the problem which was the subject of counsel's email.

In the ensuing days, I learned that the database in both Langford and Coley had been produced using an additive process, meaning that, instead of starting with the entire database and taking out what was not to be produced, they constructed "queries" in which they attempted to define what records to produce. In the subtractive process, the default is that every record is produced, except what is specifically agreed will not be produced. In the additive process, the default is that nothing will be produced unless it is specifically identified.

This was directly contrary to what was agreed upon in Langford. The original dispute, evidenced in Exhibit 5, arose from defendants' intention to produce a subset of the database which it believed related to plaintiffs' efforts to obtain work. That dispute was resolved by the

Union's agreement to produce the entire database, minus some personal information relating to members. This understanding was further expressed in an email from counsel for defendant dated April 20, 2021, annexed hereto as Exhibit 7, in which she stated in reference to the referral database that defendant "provided *everything* except [enumerated personal information]."

The problem described by counsel for defendant on October 17, 2023 arose from defendant using different queries to identify data to be produced in Langford and this case. These queries were identified in an email dated October 20, 2023, a copy of which is annexed hereto as Exhibit 8. The problem with using an additive process is demonstrated by this snafu. Both query authors intended (presumably) to pull the same records, but ended up with different data sets.

Defendant's database consisted of several dozen tables. The error revealed in counsel's October 17, 2023 email related to just one of those tables. I have requested defendant to provide us with the queries used for all of the tables, both in Langford and this case. I have also requested defendant to produce a witness to answer questions about the queries. Defendant has refused.

The requested discovery is necessary because the query for the job referral table revealed that potentially hundreds of relevant records have been omitted from production.

The accompanying declaration of Shane Thompson, PhD, demonstrates that the query used to pull records in both Langford and this case was created in such a way as to exclude records that should have been produced. The problem is that the query required two temporal conditions to be met, even though a single temporal condition was sufficient to require inclusion of the record.

The temporal conditions were specified for two data fields: one that registered the date of the referral, and the other that registered the date that the referral was entered into the computer system. One would naturally expect both to be the same date, except perhaps in cases where a referral was made after hours for the next day. Inexplicably, however, hundreds of referrals were entered into the computer system more than 30 days before the actual job referral. Putting aside the curiosity of this fact, what it means is that the requirement that the date of *both* fields be within the relevant time period guarantees that numerous referral records will be excluded.

This fact was unknown and unknowable by plaintiff until we were shown the query used to pull records. Up to that point, we were working under the assumption that *all* referral records within the relevant time period were being produced, as explicitly represented by counsel for defendant in her April 21, 2021 email (exhibit 7).

I will note that this is not the first time that defendant's production of its database has proven to exclude relevant records. In *Langford*, the job referral database was re-produced twice, due to omissions. The first time, defendant failed to produce job referrals of non-

members who make use of the referral hall[1], and the second time was when we discovered that the job referral records were missing data necessary to determine a member's place on the out of work list. These deficiencies were discovered in the course of our analysis of defendant's database, but the deficiency uncovered here did not and would not reveal itself in the course of normal analysis.

The requested discovery is necessary. We need to see the queries for all the database tables both to ensure that the data in Langford and in Coley were both produced the same way. We need to depose a person with knowledge about the queries to answer any questions regarding their construction, and any records that should have been produced, but which were not produced, should be produced.

Plaintiff has demonstrated "good cause."

Reopening discovery is within the sound discretion of the Court, and it should be granted where there is a showing of "good cause." *Oliver v. Am. Express Co.*, 2022 U.S. Dist. LEXIS 239881, *12 (E.D.N.Y. 2022). There is no question that where a party produces relevant information after the close of discovery there is good cause to reopen discovery. *Scantibodies Lab., Inc. v. Church & Dwight Co.*, 2016 U.S. Dist. LEXIS 154396, *131 (S.D.N.Y. 2016). Plaintiff has acted diligently in moving to reopen discovery upon learning of the problems with defendant's initial production of its database.

If discovery is reopened, defendant should be required to provide year of birth data for its members.

Earlier this year, the Court denied plaintiff's request to compel defendant to produce the year of birth information for the members in its database. See the Court's June 13, 2023 docket entry. The Court granted defendant's application for a protective order, which was based on the assertion that (a) the request was untimely and (b) defendant's production of its database was "not incomplete or incorrect in any material respect." Given the developments discussed herein, the basis for the protective order no longer exists.

Respectfully yours,

---

1   These include trainees and individuals working on "permit" or "clearance."